Probability Comprehensive Exam
Fall 2018

Student Number:

Instructions: Complete 5 of the 9 problems, and **circle** their numbers below – the uncircled problems will **not** be graded.

1 2 3 4 5 6 7 8 9

Write **only on the front side** of the solution pages. A **complete solution** of a problem is preferable to partial progress on several problems.

1. Use the SLLN to find the following limit:

$$\lim_{n\to\infty}\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n.$$

Solution: Let $U_1,\ldots,U_n$ be i.i.d. random variables with uniform distribution in $[0,1]$. Then

$$\mathbb{E}\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}=\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n.$$

By the SLLN,

$$\frac{U_1^2+\cdots+U_n^2}{n}\to \mathbb{E}U_1^2=\int_0^1 x^2dx=\frac{1}{3}\text{ a.s.}$$

and

$$\frac{U_1+\cdots+U_n}{n}\to \mathbb{E}U_1=\int_0^1 xdx=\frac{1}{2}\text{ a.s.}$$

Therefore,

$$\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}=\frac{(U_1^2+\cdots+U_n^2)/n}{(U_1+\cdots+U_n)/n}\to\frac{2}{3}\text{ a.s.}$$

Since also

$$0\le \frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}\le 1,$$

we have, by Lebesgue dominated convergence, that

$$\int_0^1\cdots\int_0^1 \frac{x_1^2+\cdots+x_n^2}{x_1+\cdots+x_n}dx_1\ldots dx_n=\mathbb{E}\frac{U_1^2+\cdots+U_n^2}{U_1+\cdots+U_n}\to\frac{2}{3}$$

as $n\to\infty$.

2. Suppose $X_1,\ldots,X_n$ are i.i.d. random variables such that $\mathbb{P}\{X_j=+1\}=\mathbb{P}\{X_j=-1\}=1/2$. Let $S_k:=X_1+\cdots+X_k, k=1,\ldots,n$. Prove that

$$\mathbb{P}\{\max_{1\le k\le n}S_k\ge l\}=2\mathbb{P}\{S_n>l\}+\mathbb{P}\{S_n=l\}.$$

Solution: Note that, by additivity and independence,

$$\begin{aligned}\mathbb{P}\{S_n > l\} &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l, S_n - S_k > 0\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}\mathbb{P}\{S_n - S_k > 0\}\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}\{S_n = l\} &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l, S_n - S_k = 0\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}\mathbb{P}\{S_n - S_k = 0\}.\end{aligned}$$

This implies that

$$\begin{aligned}&2\mathbb{P}\{S_n > l\} + \mathbb{P}\{S_n = l\} \\ &= \sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\}(2\mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k = 0\}).\end{aligned}$$

Finally, note that by symmetry of r.v. $S_n - S_k = \sum_{j=k+1}^{n} X_j$,

$$\begin{aligned}&2\mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k = 0\} \\ &= \mathbb{P}\{S_n - S_k > 0\} + \mathbb{P}\{S_n - S_k < 0\} + \mathbb{P}\{S_n - S_k = 0\} = 1\end{aligned}$$

and

$$\sum_{k=1}^{n} \mathbb{P}\{S_1 < l, \ldots, S_{k-1} < l, S_k = l\} = \mathbb{P}\{\max_{1 \le k \le n} S_k \ge l\},$$

implying the claim.

3. Let $\{Z_n\}$ be i.i.d. standard normal r.v. and let $\{a_n\}$ be a sequence of nonnegative real numbers. Prove that $\sum_{n=1}^{\infty} a_n Z_n^2 < +\infty$ a.s. if and only if $\sum_{n=1}^{\infty} a_n < +\infty$.

Solution: If $\sum_{n=1}^{\infty} a_n < +\infty$, then

$$\mathbb{E}\sum_{n=1}^{\infty} a_n Z_n^2 = \sum_{n=1}^{\infty} a_n \mathbb{E} Z_n^2 = \sum_{n=1}^{\infty} a_n < +\infty,$$

implying that the nonnegative r.v. $\xi := \sum_{n=1}^{\infty} a_n Z_n^2$ is finite a.s. On the other hand, if $\xi < +\infty$ a.s., then $e^{-\xi} > 0$ a.s., implying that $\mathbb{E}e^{-\xi} > 0$. By a straightforward computation,

$$\mathbb{E}e^{-\xi} = \prod_{n=1}^{\infty} \mathbb{E}e^{-a_n Z_n^2} = \prod_{n=1}^{\infty} \mathbb{E}e^{-a_n Z_1^2} = \prod_{n=1}^{\infty} \frac{1}{\sqrt{1+2a_n}}$$

The last product is strictly positive if and only if the series $\sum_{n=1}^{\infty} \log(1+2a_n)$ converges, which implies $\sum_{n=1}^{\infty} a_n < +\infty$.

4. Let φ be the characteristic function of r.v. X. Show that

$$\psi_1(t) = |\varphi(t)|^2 \text{ and } \psi_2(t) = \frac{1}{t}\int_0^t \varphi(s)ds$$

are also characteristic functions.

Solution: Note that

$$\psi_1(t) = \varphi(t)\overline{\varphi(t)} = \mathbb{E}e^{itX}\mathbb{E}e^{-itX} = \mathbb{E}e^{itX}\mathbb{E}e^{-itY} = \mathbb{E}e^{it(X-Y)},$$

where Y is an independent copy of X. Thus, ψ_1 is the characteristic function of $X-Y$.

By change of variable and the properties of conditional expectation,

$$\psi_2(t) = \frac{1}{t}\int_0^t \varphi(s)ds = \int_0^1 \varphi(tu)du = \int_0^1 \mathbb{E}e^{itXu}du$$
$$= \int_0^1 \mathbb{E}(e^{itXU}|U=u)du = \mathbb{E}\mathbb{E}(e^{itXU}|U) = \mathbb{E}e^{itXU},$$

where U is a random variable with uniform distribution in $[0,1]$ independent of X. Thus, ψ_2 is the characteristic function of XU.

5. For distribution functions F, G on the real line, define

$$L(F,G) := \inf\Big\{\varepsilon > 0 : \forall t \in \mathbb{R}\ F(t) \le G(t+\varepsilon)+\varepsilon, G(t) \le F(t+\varepsilon)+\varepsilon\Big\}.$$

It is known that L is a metric. Prove that $L(F_n, F) \to 0$ as $n \to \infty$ if and only if F_n converges weakly to F.

Solution: If $L(F_n, F) \to 0$ as $n \to \infty$, then, for any $\varepsilon > 0$ and all large enough n, $L(F_n, F) < \varepsilon$. This implies that, for all large enough n,

$$\forall t \; F(t-\varepsilon) - \varepsilon \le F_n(t) \le F(t+\varepsilon) + \varepsilon.$$

Therefore

$$F(t-\varepsilon) - \varepsilon \le \liminf_{n\to\infty} F_n(t) \le \limsup_{n\to\infty} F_n(t) \le F(t+\varepsilon) + \varepsilon. \tag{1}$$

Passing to the limit when $\varepsilon \to 0$, we get

$$F(t-) \le \liminf_{n\to\infty} F_n(t) \le \limsup_{n\to\infty} F_n(t) \le F(t). \tag{2}$$

If t is a continuity point of F, we have $F(t) = F(t-)$ and

$$\lim_{n\to\infty} F_n(t) = F(t),$$

which implies the weak convergence of F_n to F.

On the other hand, the weak convergence of F_n to F easily implies (2), which implies (1). It follows from (1) and the definition of L that $L(F_n, F) < 2\varepsilon$ for all n large enough. Therefore, $L(F_n, F) \to 0$ as $n \to \infty$.

6. Let $X_1, X_2, \ldots, X_n, \ldots$ be identically distributed (not necessarily independent!) random variables with finite first moment. Is the following,

$$n^{-1}\mathbb{E} \max_{1\le k\le n} |X_k| \longrightarrow 0,$$

as $n \to +\infty$, true or false?

Solution: True! Indeed, for any $A > 0$, and using the identical distribution assumption,

$$
\begin{aligned}
\mathbb{E}\max_{1\le k\le n}|X_k| &= \int_0^{+\infty}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt \\
&= \int_0^{A}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt + \int_A^{+\infty}\mathbb{P}(\max_{1\le k\le n}|X_k|>t)dt \\
&\le A + \int_A^{+\infty}\sum_{k=1}^{n}\mathbb{P}(|X_k|>t)dt \\
&= A + n\int_A^{+\infty}\mathbb{P}(|X_1|>t)dt.
\end{aligned}
$$

Therefore, for any $A>0$,

$$\limsup_{n\to+\infty}\frac{1}{n}\mathbb{E}\max_{1\le k\le n}|X_k| \le \int_A^{+\infty}\mathbb{P}(|X_1|>t)dt.$$

But, $\mathbb{E}|X_1| = \int_0^{+\infty}\mathbb{P}(|X_1|>t)dt < +\infty$, and so by dominated convergence,

$$\limsup_{A\to+\infty}\limsup_{n\to+\infty}\frac{1}{n}\mathbb{E}\max_{1\le k\le n}|X_k| \le 0,$$

which gives the result.

7. Let $X_1, X_2, \dots, X_n, \dots$ be iid random variables with common characteristic function φ and let $S_n = \sum_{k=1}^n X_k$. Show that if φ is differentiable at 0 with $\varphi'(0) = i\mu$, then, as $n\to+\infty$, $S_n/n \to \mu$, in probability.

Solution: In case the limit is degenerate then convergence in probability is equivalent to weak convergence. In other words, $S_n/n \to \mu$, in probability if and only if $S_n/n \Rightarrow \mu$. In turn by the Lévy continuity theorem, this last condition is equivalent to the requirement that for all $t\in\mathbb{R}$, $\mathbb{E}(e^{itS_n/n}) \to e^{it\mu}$. Now by the iid assumption, $\mathbb{E}(e^{itS_n/n}) = (\varphi(t/n))^n$. Since φ is differentiable at 0,

$$\lim_{n\to+\infty}\frac{\varphi(t/n)-1}{t/n} = \varphi'(0) = i\mu,$$

i.e., $\lim_{n\to+\infty} n(\varphi(t/n)-1) = it\mu$. Finally, since

$$(\varphi(t/n))^n = \left(1+\frac{n(\varphi(t/n)-1)}{n}\right)^n,$$

using complex logarithms or the fact that if $z_n\in\mathbb{C}$ is such that $z_n\to z\in\mathbb{C}$, then $(1+z_n/n)^n \to e^z$, the result follows.

8. Let X and Y be two independent and positive random variables with respective density f_X and f_Y and let $g : (0, +\infty) \longrightarrow (0, +\infty)$, be a bounded Borel function. Find

$$\mathbb{E}\left(g\left(\frac{X}{Y}\right)|Y\right),$$

the conditional expectation of $g(X/Y)$ given Y and then infer that $V = X/Y$ has a density that you will identify.

Solution: Since X and Y are independent, $\mathbb{E}\left(g\left(\frac{X}{Y}\right)|Y\right) = h(Y)$, with $h(y) = \mathbb{E}\left(g(X/y)\right)$. Therefore,

$$\begin{aligned} h(y) &= \int_0^{+\infty} g\left(\frac{x}{y}\right) f_X(x)dx \\ &= y\int_0^{+\infty} g(v) f_X(yv)dv. \end{aligned}$$

Next, for any g as above,

$$\mathbb{E}g(V) = \mathbb{E}(\mathbb{E}(g(V)|Y)) = \mathbb{E}h(Y).$$

But, using the Fubini-Tonelli Theorem which is valid since all our functions are non-negative as well as Lebesgue measurable,

$$\begin{aligned} \mathbb{E}g(V) = \mathbb{E}h(Y) &= \int_0^{+\infty} h(y) f_Y(y)dy \\ &= \int_0^{+\infty} f_Y(y)\left(\int_0^{+\infty} yg(v) f_X(yv)dv\right)dy \\ &= \int_0^{+\infty} g(v)\left(\int_0^{+\infty} yf_Y(y) f_X(yv)dy\right)dv \\ &= \int_0^{+\infty} g(v) f(v)dv, \end{aligned}$$

where $f(v) := \int_0^{+\infty} yf_Y(y)f_X(yv)dy$ is therefore the density of V.

9. Let X, Y, Z be random variables such that (X, Z) and (Y, Z) are identically distributed. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a Borel function such that $f(X)$ is integrable.
(i) Show that $\mathbb{E}(f(X)|Z) = \mathbb{E}(f(Y)|Z), a.s.$
(ii) Let $T_1, T_2, \ldots T_n$ be iid random variables with finite first moment and let $T = T_1 + \cdots + T_n$. Using (i) show that

$$\mathbb{E}(T_1|T) = \frac{T}{n}.$$

Solution: (i) For any non-negative (or bounded) Borel function g, since $g(Z)$ is Z-measurable, since the "expectation of the conditional expectation is the expectation", and using the identical distribution assumption,

$$\begin{aligned}\mathbb{E}(g(Z)\mathbb{E}(f(X)|Z)) &= \mathbb{E}(\mathbb{E}(g(Z)f(X)|Z)) = \mathbb{E}(g(Z)f(X)) \\ &= \mathbb{E}(g(Z)f(Y)) = \mathbb{E}(\mathbb{E}(g(Z)f(Y)|Z)) \\ &= \mathbb{E}(g(Z)\mathbb{E}(f(Y)|Z)),\end{aligned}$$

from which it follows (by the very definition and uniqueness of the conditional expectation) that $\mathbb{E}(f(X)|Z) = \mathbb{E}(f(Y)|Z), a.s.$, since both quantities above are $Z-$measurable.

(ii) Clearly, $(T_1, T), (T_2, T), \ldots, (T_n, T)$ are identically distributed and so, by (i),

$$\mathbb{E}(T_1|T) = \mathbb{E}(T_2|T) = \cdots = \mathbb{E}(T_n|T).$$

Therefore,

$$\begin{aligned}n\mathbb{E}(T_1|T) &= \mathbb{E}(T_1|T) + \mathbb{E}(T_2|T) + \cdots + \mathbb{E}(T_n|T) \\ &= \mathbb{E}(T_1 + T_2 + \cdots + T_n|T) \\ &= \mathbb{E}(T|T) = T,\end{aligned}$$

which shows that $\mathbb{E}(T_1|T) = T/n$.